A Direct Sum Theorem for Corruption and the Multiparty NOF Communication Complexity of Set Disjointness

Paul Beame*
University of Washington
Seattle, WA 98195-2350
beame@cs.washington.edu

Toniann Pitassi†
University of Toronto
Toronto, ON M5S 1A4
toni@cs.toronto.edu

Nathan Segerlind‡
University of Washington
Seattle, WA 98195-2350
nsegerli@cs.washington.edu

Avi Wigderson§
Institute for Advanced Study
Princeton, NJ
avi@ias.edu

Abstract

We prove that corruption*, one of the most powerful measures used to analyze 2-party randomized communication complexity, satisfies a strong direct sum property under rectangular distributions. This direct sum bound holds even when the error is allowed to be exponentially close to 1. We use this to analyze the complexity of the widely-studied set disjointness problem in the usual "number-on-the-forehead" (NOF) model of multiparty communication complexity.*

1 Introduction

One of the most important problems in communication complexity is the two party set disjointness function: Alice and Bob are each given a subset of $[n]$ and they want to determine whether or not they share a common element [2, 17, 28, 27]. A natural extension of two party disjointness is k party disjointness. Now there are k players, each with private inputs $x_1, \ldots, x_k \subseteq [n]$ respectively, and again they want to determine whether or not they share a common element. This problem, multiparty set disjointness in the "number-in-hand" (NIH) model, has been the focus of considerable research largely because randomized lower bounds in this setting are related to the space complexity of randomized streaming algorithms that approximately compute frequency moments of a data set [1]. After a sequence of improved bounds [1, 29, 6, 7], nearly optimal bounds are now known [10] for multiparty NIH set disjointness.

Another model of multiparty communication complexity that is frequently studied is the "number-on-the-forehead" (NOF) model [12] in which each player sees all but his/her own input which metaphorically is on his/her forehead. This model is extremely important, as linear lower bounds for $k = n^\epsilon$ players for any explicit function would solve a major open problem in circuit complexity. The only lower bounds known so far hold for $k \leq \log_2 n$ players, and, with one exception, use the discrepancy method [5, 13, 26] in which it is required to show that the function is nearly balanced on all large cylinder intersections. Thus, it is a major challenge to develop new techniques for understanding NOF multi-party protocols. Interestingly, the multiparty NOF complexity of set disjointness even for three players is almost completely open, and seems to be an important step towards understanding multiparty NOF protocols in general. This is because the discrepancy technique yields only trivial bounds for set disjointness (because set disjointness is constant on some very large cylinder intersections), so progress here should involve a new kind of argument.

Additionally there are several other interesting applications of strong lower bounds for multiparty NOF complexity of set disjointness. By a natural extension of the ideas in [2] one can show that for $k \geq 2$ the k-party set disjointness problem is "complete" for the k-party communication complexity class k-NP^{cc}. Thus, strong lower bounds for set disjointness would prove a separation between nondeterministic and deterministic (or randomized) multiparty NOF communication complexity. Secondly, as we show in re-

*Supported by NSF grants CCR-0098066 and ITR-0219468
†Supported by an Ontario Premiere's Research Excellence Award, an NSERC grant, and the Institute for Advanced Study. Research done while at the Institute for Advanced Study.
‡Supported by NSF Postdoctoral Fellowship DMS-0303258
§Partially supported by NSF grant CCR-0324906

Proceedings of the Twentieth Annual IEEE Conference on Computational Complexity (CCC'05)
1093-0159/05 $20.00 © 2005 IEEE

lated work [8], $\omega(\log^3 n)$ lower bounds for k-party NOF set disjointness yield lower bounds for a large family of proof systems known as tree-like, degree $k-1$ threshold systems. Such systems are quite powerful, and include systems such as the Chvatal-Gomory Cutting Planes proof system, and all variations of the Lovasz-Schriver proof systems, in tree-form. (We further showed that lower bounds of the form $\omega(\log n(\log\log n)^2)$ also have non-trivial proof complexity consequences.)

The best protocol known for the k-party NOF set disjointness problem is the trivial one with complexity n in which one player broadcasts its input; given the bounds for the two-party case, when k is constant one should expect the complexity of set disjointness to be $n^{\Omega(1)}$ (if not $\Omega(n)$). The only previous lower bound for k-party NOF set disjointness for $k \geq 3$ is a bound of Wigderson that appears in [4] which shows that the one-way 3-party NOF complexity of set disjointness is $\Omega(\sqrt{n})$. (The bound as stated is for a layered pointer jumping problem which corresponds to the special case of the disjointness problem in which the first player's input is one of $\sqrt{n}$ disjoint subsets of $[n]$ of size $\sqrt{n}$, the second player's input has one element in each of these $\sqrt{n}$ blocks and the third player's input is an arbitrary vector of n bits.)

In this paper, we make some further progress toward understanding the NOF multiparty complexity of set disjointness, by deriving several new lower bounds. We prove that in the *simultaneous* NOF communication model [3] the randomized k-party complexity of disjointness is $\Omega(n^{1/(k-1)}/(k-1))$. We also obtain an $\Omega(n^{1/3})$ lower bound in a model of 3-party NOF communication complexity that does not have a one-way or simultaneous requirement – in particular in a model (implicitly considered by Nisan and Wigderson [24]) in which one of the three parties communicates once at the start and then the other two parties alternate communication arbitrarily. Finally, in the general model needed for the proof complexity bounds above, we obtain an $\Omega(\log n)$ lower bound for randomized 3-party NOF communication complexity of set disjointness using techniques related to our $\Omega(n^{1/3})$ lower bound and a $\Omega(\frac{\log n}{k-1})$ lower bound for its randomized k-party communication complexity based on our lower bound for simultaneous protocols.

To date, other than the Ramsey-theoretic bounds shown in [12], general NOF communication complexity lower bounds have all been proved using discrepancy which does not suffice for analyzing disjointness. Our bounds are particularly interesting because they introduce a new method for proving multiparty NOF bounds via an extension of a measure used to analyze 2-party communication complexity. As a key part of our argument we show that this 2-party measure satisfies a direct sum property. This result is interesting in its own right.

For a function $f : I \to O$, the function $f^t : I^t \to O^t$ given by $f^t(x_1, \ldots, x_t) = (f(x_1), \ldots, f(x_t))$. A complexity measure C satisfies a *direct sum property* if and only if $C(f^t) = \Omega(tC(f))$. This property is *strict* if and only if the lower bound is essentially $tC(f)$ rather than $\Omega(tC(f))$. Karchmer, Raz, and Wigderson [19] introduced the direct sum problem in 2-party communication complexity (in particular of Boolean relations) and showed that its proof would yield a separation between NC^1 and NC^2.

Although the problem for Boolean relations is wide open, there are a number of direct sum results for 2-party communication complexity of functions which have been shown to have independent interest. A strict direct sum property is known for nondeterministic and co-nondeterministic 2-party communication complexity and direct sum properties are known for bounded-round deterministic [18, 15] and bounded-round distributional/randomized [16] 2-party communication complexity. (Note that one should only consider randomized computation with *public* randomness since the direct sum property is not strictly true with private randomness, although there is no essential distinction between these at complexities that are $\Omega(\log n)$. Note also that an alternative way to express Wigderson's bound for one-way 3-party NOF set disjointness in [4] can be derived from direct sum results for randomized one-way 2-party communication complexity.)

While there are no universal direct sum results for unrestricted deterministic or randomized 2-party communication complexity, there are a number of useful direct sum results for the measures used to derive lower bounds for these complexities. The first of these results was the observation [19] that the rank of the communication matrix used for deterministic communication complexity lower bounds satisfies a strict direct sum property.

For randomized communication complexity, Shaltiel [30] proved that one major measure, discrepancy, satisfies a direct product property and therefore satisfies a direct sum property. Moreover, one of the most important values of information complexity [11, 6] and conditional information complexity [7] as lower bound techniques for randomized communication complexity is the very fact that these measures satisfy direct sum properties under rectangular (or conditionally rectangular) distributions.

Information complexity and discrepancy are incomparable measures [6]. We show that a strictly more general measure than discrepancy, *corruption*, which is often a measure of choice for randomized communication complexity bounds, satisfies a direct sum property for rectangular distributions.[1] Corruption bounds, which we define in

[1] Although corruption bounds are frequently used, there does not seem to be a consistent terminology for such bounds. We hope that the term "corruption" is a suitable complement to "discrepancy".

Proceedings of the Twentieth Annual IEEE Conference on Computational Complexity (CCC'05)
1093-0159/05 $20.00 © 2005 IEEE

the next section, were used for example lower bounds for the 2-party communication complexity of the set disjointness function [2, 28], for which discrepancy is ineffective. Moreover, Klauck [20] has shown that bounds based on corruption can be exponentially better than those based on discrepancy.

We use the direct sum property for corruption to analyze the direct sum of the set disjointness function and apply the results in analyzing set disjointness in the multiparty NOF model.

2 Discrepancy, Corruption, and Communication Complexity

Let $f : I \to O$. For $b \in O$, a subset $S \subseteq I$ is called *b-monochromatic* for f if and only if $f(s) = b$ for all $s \in S$ and is called *monochromatic* if and only if it is b-monochromatic for f for some $b \in O$.

Let μ be a probability measure on I. For $b \in O$, a subset $S \subseteq I$ is called *ϵ-error b-monochromatic* for f under μ if and only if $\mu(S \cap \overline{f^{-1}(b)}) \leq \epsilon \cdot \mu(S)$. For $f : I \to \{0,1\}$, $b \in \{0,1\}$, and $S \subseteq I$ the *b-discrepancy* of f on S under μ,

$$\mathrm{disc}^b_\mu(f, S) = \mu(S \cap f^{-1}(b)) - \mu(S \cap f^{-1}(\overline{b})).$$

Let Γ be a collection of subsets of I. Writing monoχ for monochromatic define

$$\mathrm{mono}^b_{\mu,\Gamma}(f) = \max\{\mu(S) \mid S \in \Gamma \text{ is } b\text{-mono}\chi\},$$
$$\epsilon\text{-mono}^b_{\mu,\Gamma}(f) = \max\{\mu(S) \mid S \in \Gamma \text{ is } \epsilon\text{-error } b\text{-mono}\chi\},$$
$$\mathrm{disc}^b_{\mu,\Gamma}(f) = \max\{\mathrm{disc}^b_\mu(f, S) \mid S \in \Gamma\}.$$

For $\Gamma \subseteq \mathcal{P}(I)$, define $\mathrm{mono}_{\mu,\Gamma}(f) = \max\{\mathrm{mono}^b_{\mu,\Gamma}(f) \mid b \in O\}$ and for $f : I \to \{0,1\}$ the *discrepancy* of f under μ, $\mathrm{disc}_{\mu,\Gamma}(f) = \max\{\mathrm{disc}^0_{\mu,\Gamma}(f), \mathrm{disc}^1_{\mu,\Gamma}(f)\}$; when μ is omitted it is assumed to be the uniform distribution.

The following is a simple relationship between these measures.

Proposition 1. *For any function $f : I \to \{0,1\}$, distribution μ on I, $\Gamma \subseteq \mathcal{P}(I)$, $\epsilon < 1/2$, and $b \in \{0,1\}$, $\mathrm{disc}^b_{\mu,\Gamma}(f) \geq (1 - 2\epsilon)[\epsilon\text{-mono}^b_{\mu,\Gamma}(f)]$.*

Proof. Let $S \in \Gamma$ witness the value of $\epsilon\text{-mono}^b_{\mu,\Gamma}(f)$ so that $\mu(S) = \epsilon\text{-mono}^b_{\mu,\Gamma}(f)$ and $\mu(S \cap f^{-1}(\overline{b})) \leq \epsilon\mu(S)$. Then $\mathrm{disc}^b_{\mu,\Gamma}(f) \geq \mathrm{disc}^b_\mu(f, S) \geq (1 - 2\epsilon)\mu(S)$ as required. □

A *combinatorial rectangle* R on set $X \times Y$ is an element $A \times B$ of $\mathcal{P}(X) \times \mathcal{P}(Y)$. Combinatorial rectangles are naturally associated with 2-party communication complexity. Let $D^2(f)$ ($N^2_1(f)$, $N^2_0(f)$) be the 2-party deterministic (respectively nondeterministic, co-nondeterministic) communication complexity of a function $f : X \times Y \to O$. For example, the following is a standard way to obtain communication complexity lower bounds.

Proposition 2. *Let Γ be the set of combinatorial rectangles on $X \times Y$. For any $f : X \times Y \to \{0,1\}$ and for any probability measure μ on $X \times Y$,*

(a) $D^2(f) \geq \log_2(1/\mathrm{mono}_{\mu,\Gamma}(f))$,

(b) For $b \in \{0,1\}$, $N^2_b(f) \geq \log_2(\mu(f^{-1}(b))/\mathrm{mono}^b_{\mu,\Gamma}(f))$.

An *i-cylinder* C on $U = X_1 \times \cdots \times X_k$ is a set of the form

$$\{(x_1, \ldots, x_k) \in U \mid g(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k) = 1\}$$

for some function g. A *cylinder intersection* on $X_1 \times \cdots \times X_k$ is a set $E = \bigcap_{i=1}^k C_i$ where C_i is an i-cylinder on $X_1 \times \cdots \times X_k$. Cylinder intersections are naturally associated with the so-called "number-on-the forehead" (NOF) model of k-party communication complexity in which the i-th player sees every portion of the input except for the i-th portion. For $k = 2$, cylinder intersections are combinatorial rectangles.

Let $R^k_\epsilon(f)$ be the ϵ-error k-party randomized NOF communication complexity of f. The following are the standard discrepancy lower bounds for randomized communication complexity (see for example [22]).

Proposition 3 (Discrepancy Bound). *Let Γ be the set of combinatorial rectangles on $X \times Y$. Let $f : X \times Y \to \{0,1\}$, $\epsilon < 1/2$, and μ be any probability distribution on $X \times Y$.*

(a) $R^2_\epsilon(f) \geq \log_2((1 - 2\epsilon)/\mathrm{disc}_{\mu,\Gamma}(f))$;

(b) for $b \in \{0,1\}$, $R^2_\epsilon(f) \geq \log_2[(\mu(f^{-1}(b)) - \epsilon)/\mathrm{disc}^b_{\mu,\Gamma}(f)]$.

More generally, for $k \geq 2$, if $f : X_1 \times \cdots \times X_k \to \{0,1\}$ and Γ is replaced by the set of cylinder intersections on $X_1 \times \cdots \times X_k$ then

$$R^k_\epsilon(f) \geq \log_2[(\mu(f^{-1}(b)) - \epsilon)/\mathrm{disc}^b_{\mu,\Gamma}(f)].$$

The $1 - 2\epsilon$ in the numerator in (a) is the required total correlation of the output of the protocol with the function f. The numerator in (b) is the required correlation of the b outputs of the protocol with the value of f. Observe that the bound from part (a) can never be more than 1 more than the maximum of the two bounds from part (b).

The discrepancy bound works well for functions such as the inner product function in two party communication complexity and for generalized inner product for multiparty communication complexity.

However, it does not suffice to derive lower bounds for the set disjointness problem among others. An alternative

lower bound method that suffices for 2-party lower bounds for set disjointness is based on a *corruption bound* which says that any sufficiently large rectangle cannot be fully b-monochromatic and is thus "corrupted" by some fixed fraction of errors. Although it is implicit in many lower bound arguments we state it explicitly here.

Lemma 4 (Corruption Bound). *Let Γ be the set of combinatorial rectangles on $X \times Y$. Let $f : X \times Y \to O$, $O' \subset O$, $\epsilon \leq 1$, and μ be any probability distribution on $X \times Y$. For $\epsilon' < \epsilon \cdot \mu(f^{-1}(O'))$,*

$$R^2_{\epsilon'}(f) \geq \min_{b \in O'} \log_2[(\mu(f^{-1}(O')) - \epsilon'/\epsilon)/\epsilon\text{-mono}^b_{\mu,\Gamma}(f))].$$

More generally, for $k \geq 2$, if $f : X_1 \times \cdots \times X_k \to O$ and Γ is replaced by the set of cylinder intersections on $X_1 \times \cdots \times X_k$ then

$$R^k_{\epsilon'}(f) \geq \min_{b \in O'} \log_2[(\mu(f^{-1}(O')) - \epsilon'/\epsilon)/\epsilon\text{-mono}^b_{\mu,\Gamma}(f))].$$

Proof. We give the proof for $k = 2$; the argument for $k > 2$ is completely analogous. By Yao's Lemma $R^2_{\epsilon'}(f)$ is at least the maximum number of bits communicated by the best ϵ'-error deterministic protocol under distribution μ. Fix such any deterministic protocol and consider the partition of $X \times Y$ into rectangles induced by the protocol. Let $\gamma = \max_{b \in O'} \epsilon\text{-mono}^b_{\mu,\Gamma}(f)$. For $b \in O'$, let α_b be the total measure of inputs contained in rectangles of measure at most γ on which the protocol outputs b. There must be at least $\sum_{b \in O'} \alpha_b/\gamma$ such rectangles and thus $R^2_{\epsilon'}(f) \geq \log_2(\sum_{b \in O'} \alpha_b/\gamma)$.

We now bound $\sum_{b \in O'} \alpha_b$. For any $b \neq b' \in O$, let $\epsilon'_{b \to b'}$ the total measure of inputs on which the protocol answers b' when the correct answer is b. Clearly $\epsilon' = \sum_{b,b' : b \neq b'} \epsilon'_{b \to b'}$. By definition, the protocol answers b on at least a $\mu(f^{-1}(b)) + \sum_{b' \neq b} \epsilon'_{b' \to b} - \sum_{b' \neq b} \epsilon'_{b \to b'}$ measure of the inputs in $f^{-1}(b)$. By the definition of γ and $\epsilon\text{-mono}^b_{\mu,\Gamma}(f)$, any rectangle of measure larger than γ on which the protocol answers b must have at least an ϵ proportion of its total measure on which the correct answer is not b; i.e., an ϵ proportion of its measure contributes to $\sum_{b' \neq b} \epsilon'_{b' \to b}$. Thus in total for $b \in O$ we have

$$\sum_{b' \neq b} \epsilon'_{b' \to b} \geq \epsilon \cdot [\mu(f^{-1}(b)) + \sum_{b' \neq b} \epsilon'_{b' \to b} - \alpha_b - \sum_{b' \neq b} \epsilon'_{b \to b'}].$$

Rearranging, we have

$$\alpha_b \geq \mu(f^{-1}(b)) - \sum_{b' \neq b} \epsilon'_{b \to b'} - (1/\epsilon - 1) \sum_{b' \neq b} \epsilon'_{b' \to b}.$$

Summing this over all choices of $b \in O'$ we obtain

$$\begin{aligned}
\sum_{b \in O'} \alpha_b &\geq \sum_{b \in O'} \mu(f^{-1}(b)) - \sum_{b \in O'} \sum_{b' \neq b} \epsilon'_{b \to b'} \\
&\quad - (1/\epsilon - 1) \sum_{b \in O'} \sum_{b' \neq b} \epsilon'_{b' \to b} \\
&= \mu(f^{-1}(O')) - (1/\epsilon) \sum_{b,b' \in O' : b \neq b'} \epsilon'_{b \to b'} \\
&\quad - \sum_{b \in O'} \sum_{b' \notin O'} \epsilon'_{b \to b'} - (1/\epsilon - 1) \sum_{b \notin O'} \sum_{b' \in O'} \epsilon'_{b \to b'} \\
&\geq \mu(f^{-1}(O')) - (1/\epsilon) \sum_{b,b' : b \neq b'} \epsilon'_{b \to b'} \\
&= \mu(f^{-1}(O')) - \epsilon'/\epsilon
\end{aligned}$$

which yields the claimed lower bound. □

In the special case that the output set $O = \{0, 1\}$ we obtain the following corollary.

Corollary 5. *Let Γ be the set of combinatorial rectangles on $X \times Y$. For any $\epsilon < 1/2$ there is a constant $c_\epsilon > 0$ such that for $f : X \times Y \to \{0, 1\}$, μ any probability distribution on $X \times Y$, and $b \in \{0, 1\}$,*

$$R^2_\epsilon(f) \geq c_\epsilon \log_2[(\mu(f^{-1}(b)) - \epsilon)/\epsilon\text{-mono}^b_{\mu,\Gamma}(f))]$$

and the same lower bound holds for the case of $R^k_\epsilon(f)$ where Γ is the corresponding set of cylinder intersections on $X_1 \times \cdots \times X_k$.

Proof. We reduce the protocol error to $\epsilon' = \epsilon^2$ and then apply Lemma 4 to obtain the claimed result. If ϵ is close to $1/2$, say $\epsilon \geq 1/3$ then repeating the ϵ-error protocol $O(1/(1 - 2\epsilon)^2)$ times and taking the majority answer ensures that the error is $\leq 1/9 \leq \epsilon^2$. If $\epsilon < 1/3$ then repeating the protocol $O(1)$ times and taking the majority answer suffices. In either case the lower bound for error ϵ^2 is reduced by only a small constant factor to yield a lower bound for error ϵ. □

Up to the constant c_ϵ, the above bound is of the same form as that of Proposition 3 except that it uses corruption rather than the discrepancy and, by Proposition 1, up to small multiplicative and additive constants, the corruption bound is superior.

3 A Direct Sum Theorem for Corruption under Rectangular Distributions

3.1 Definitions and Notation

For a function $f : X \times Y \to \{0, 1\}$, define $f^t : X^t \times Y^t \to \{0, 1\}^t$ by $f^t(\vec{x}, \vec{y}) = (f(x_1, y_1), \ldots, f(x_t, y_t))$

where $\vec{x} = (x_1, \ldots, x_t)$ and $\vec{y} = (y_1, \ldots, y_t)$. Given a distribution μ on a set I, the distribution μ^t is a distribution on I^t that is the cross product of t independent copies of μ on each of the t coordinates.

We now make some definitions for sets $S \subseteq X^t \times Y^t$. Let $T \subseteq [t]$ and $U = [t] - T$. For $S \subseteq X^t \times Y^t$, let S_T be the set of projections of S on $X^T \times Y^T$. (If T is a singleton set $\{j\}$ then we write S_j for $S_{\{j\}}$.) For $(x_U, y_U) \in X^U \times Y^U$ and $S \subseteq X^t \times Y^t$ let $S(x_U, y_U)$ be the set of all $(\vec{x}', \vec{y}') \in S$ such that $x'_U = x_U$ and $y'_U = y_U$.

A distribution μ on $X \times Y$ is *rectangular* if and only if there are probability distributions μ_X on X and μ_Y on Y such that $\mu(x, y) = \mu_X(x)\mu_Y(y)$.

Finally, we say that S is *rectangular with respect to coordinates* T if and only if for every $(x_U, y_U) \in S_U$, $S(x_U, y_U)_T$ is a combinatorial rectangle in $X^T \times Y^T$.

3.2 The Direct Sum Theorem

For $f : X \times Y \to O$ and μ a probability distribution on $X \times Y$, the *ϵ-corruption bound* for f over distribution μ with respect to b, written $\mathrm{corrbd}^b_\mu(f, \epsilon)$ is defined to be equal to $\log_2(1/\epsilon\text{-mono}^b_{\mu,\Gamma}(f))$, where Γ is the set of combinatorial rectangles on $X \times Y$. One should think of the corruption bound as a lower bound on the number of bits that need to be communicated in order to achieve a protocol for f over distribution μ with error at most ϵ.

We can re-express the bound from Corollary 5 as

$$R^2_\epsilon(f) \geq c_\epsilon(\mathrm{corrbd}^b_\mu(f, \epsilon) - \log_2(\frac{1}{\mu(f^{-1}(b) - \epsilon}))$$

and the bound from Lemma 4 as

$$R^2_{\epsilon'}(f) \geq \min_{b \in O'} \mathrm{corrbd}^b_\mu(f, \epsilon) - \log_2(\frac{1}{\mu(f^{-1}(O')) - \epsilon'/\epsilon}).$$

Theorem 6 (Direct Sum Property for Corruption). *Let $f : X \times Y \to \{0, 1\}$ and μ be a rectangular probability distribution on $X \times Y$. Let $b \in \{0, 1\}$, t be a positive integer, and let $v \in \{0, 1\}^t$ be a binary vector with at least t_0 b's. Let $m = \mathrm{corrbd}^b(f, \epsilon)$. If $1 > \epsilon > 12mt/2^{m/8}$ then $\mathrm{corrbd}^v_{\mu^t}(f^t, 1 - (3/\epsilon)(1 - \epsilon/2)^{t_0}) \geq t_0 \cdot \mathrm{corrbd}^b(f, \epsilon)/6$.*

Observe that this lemma implies very strong error properties. It says that any large rectangle on which a protocol P outputs a vector v with many b's has the correct answer on only an exponentially small fraction of the inputs under distribution μ^t.

One can loosely interpret the above theorem in the following way. Suppose that we have a lower bound of k on the number of bits that need to be communicated in order to solve one instance of f, with error at most ϵ, via a corruption bound on, say $b = 0$. (That is, we know that any rectangle for f with mostly 0's must be small.) Then the obvious protocol in order to solve t instances is to run each of the k-bit protocols in order to solve all t instances with a total of kt bits. Now, if the error is uncorrelated, the probability that the answer is correct over all of the t instances is extremely large, namely $1 - (1 - \epsilon)^t$. However, the bound will not be quite this good because the known bound for one copy/instance of f is obtained via corruption; that is, we are only guaranteed that the problem is hard on 0-instances. Thus the bit complexity of t instances as well as the error will be a function of t_0, the number of 0's in v, rather than a function of t.

As we will show shortly, we will apply the above theorem to the set disjointness problem where there is a known lower bound of $\Omega(\sqrt{n})$ for $b = 0$. This enables us to prove that solving t instances of set disjointness over a suitable distribution requires on the order of $t\sqrt{n}$ bits of communication, even to obtain a protocol that is correct on only an exponentially small fraction of the inputs.

The general technique we use for our direct sum bound follows a standard paradigm of iterated conditional probability analysis on the coordinates that allows one to prove earlier proofs of direct product theorems for circuits that allow one to prove Yao's XOR lemma [14], Raz's parallel repetition theorem [25] and bounds on the complexity savings given by 'help bits' [9, 23].

The following lemma is the main tool we need to prove the direct sum property of corruption. Its proof is the sole reason that we need to restrict the distribution μ to be rectangular.

Lemma 7 (Key Lemma). *Let $f : X \times Y \to \{0, 1\}$ and μ be a rectangular probability distribution on $X \times Y$. Let $b \in \{0, 1\}$ and $m = \mathrm{corrbd}^b_\mu(f, \epsilon)$ for $\epsilon < 1$. Let $k \geq 1$ and $A \times B \in \mathcal{P}(X^k) \times \mathcal{P}(Y^k)$. For integer $K' \geq 1$ let $K = \lceil \log_{(1-\epsilon/6)} 2^{-K'} \rceil = \lceil -K'/\log_2(1 - \epsilon/6) \rceil$. There are sets $P, Q, E \subseteq A \times B$ such that the set of inputs $(\vec{x}, \vec{y}) \in A \times B$ for which $f(x_1, y_1) = b$ is contained in $P \cup Q \cup E$ where*

- $\mu^k(E) \leq 2^{1-K'}$,
- $\mu^k(Q) \leq (1 - \epsilon/2)\mu^k(A \times B - P - E)$,
- $\mu(P_1) \leq K^2 2^{-m}$.

Furthermore P, Q, and E are rectangular on coordinates $\{2, \ldots, k\}$ and P_1, Q_1, and E_1 are all disjoint.

Proof. We would like to upper bound the fraction of inputs in $A \times B$ on which $f(x_1, y_1) = b$. The general idea of the proof involves considering the set of projections (x_1, y_1) of the elements of $A \times B$ on the first coordinate. This set forms a rectangle on $X \times Y$. By definition of $m = \mathrm{corrbd}^b_\mu(f, \epsilon)$, if this set has μ measure larger than 2^{-m} then $f(x_1, y_1) = b$ for at most a $1 - \epsilon$ fraction of the projected pairs (x_1, y_1).

However, because the different (x_1, y_1) occur with different frequencies in $A \times B$, the overall fraction of errors

may be much smaller. To overcome this problem we group the elements of A and B based on the number of extensions their projections x_1 or y_1 have in A or B respectively. We choose the groups so that each is a rectangle and in any group there is very little variation in the number of extensions. For any one of these groups containing at least a 2^{-m} fraction of (x_1, y_1) pairs we can apply the corruption bound for f to bound below 1 the fraction of inputs on which the function has output b. Any group that does not satisfy this must be small. To keep the number of groups small we first separate out one set consisting of those inputs where the number of extensions is tiny. In our argument, Q will the union of the large groups, P will be the union of the small groups, and E will be the set of inputs with a tiny number of extensions.

For $T \subseteq [k]$ define μ^T on $X^T \times Y^T$ as the cross product μ^T on those coordinates. Define μ_X^T and μ_Y^T similarly so that μ^T is the cross product of μ_X^T and μ_Y^T.

Let A_1 be the set of projections of A on the first coordinate and B_1 be the set of projections of B on the first coordinate. Choose $\delta = \epsilon/6$ and let $T = \{2, \ldots, k\}$. Sort the elements of A_1 based on the number of their extensions. For $1 \leq i \leq \lceil \log_{(1-\delta)} 2^{-K'} \rceil = \lceil -K'/\log_2(1-\delta) \rceil = K$ let $A_{1,i} = \{x_1 \in A_1 \mid i = \lceil \log_{(1-\delta)} \mu_X^T(A(x_1)_T) \rceil\}$ and $B_{1,i'} = \{y_1 \in B_1 \mid i' = \lceil \log_{(1-\delta)} \mu_Y^T(B(y_1)_T) \rceil\}$. That is, every point in $A_{1,i}$ has between a $(1-\delta)^{i-1}$ and $(1-\delta)^i$ measure of extensions in the T coordinates and similarly for each $B_{1,i'}$. Let $A^{1,i} = \{\vec{x} \in A \mid x_1 \in A_{1,i}\}$ and $B^{1,i'} = \{\vec{y} \in B \mid y_1 \in B_{1,i'}\}$. Let $E = [(A - \bigcup_{i=1}^K A^{1,i}) \times B] \cup [A \times (B - \bigcup_{i'=1}^K B^{1,i})]$. By definition $\mu^k(E) \leq 2 \cdot 2^{-K'}$.

For $i, i' \leq K$ let $R^{(i,i')} = A^{1,i} \times B^{1,i'}$ and then $A \times B = E \cup \bigcup_{i=1}^K \bigcup_{i'=1}^K R^{(i,i')}$. By definition $R_1^{(i,i')} = A_{1,i} \times B_{1,i'}$ is the projection of $R^{(i,i')}$ on the first coordinate. Also by definition, every $(x_1, y_1) \in R_1^{(i,i')}$ has at most a $(1-\delta)^{i+i'-2}$ and at least a $(1-\delta)^{i+i'}$ measure of extensions in $R^{(i,i')}$ since

$$\begin{aligned}\mu^T((R^{(i,i')}(x_1, y_1))_T) &= \mu^T(((A \times B)(x_1, y_1))_T) \\ &= \mu^T(A(x_1)_T \times B(y_1)_T) \\ &= \mu_X^T(A(x_1)_T) \cdot \mu_Y^T(B(y_1)_T)\end{aligned}$$

and for $(x_1, y_1) \in R_1^{(i,i')}$ the first quantity in the product is between $(1-\delta)^{i-1}$ and $(1-\delta)^i$ and the second is between $(1-\delta)^{i'-1}$ and $(1-\delta)^{i'}$. In particular, the measure of extensions of any two pairs $(x_1, y_1), (x_1', y_1') \in R_1^{(i,i')}$ differ by a factor between $(1-\delta)^2 \geq (1-\epsilon/3)$ and 1.

Let $G = \{(i, i') \mid \mu(R_1^{(i,i')}) = \mu(A_{1,i} \times B_{1,i'}) \geq 2^{-m}\}$. By assumption about f, for every $(i, i') \in G$,

$$\mu(A_{1,i} \times B_{1,i'} \cap f^{-1}(b)) \leq (1-\epsilon)\mu(A_{1,i} \times B_{1,i'}).$$

Let $Q^{(i,i')}$ be the set of elements of in $R^{(i,i')}$ for which the first coordinate answer is b. Since elements in $R_1^{(i,i')} = A_{1,i} \times B_{1,i'}$ have a μ^T measure of extensions in $R^{(i,i')}$ between $(1-\epsilon/3)$ and 1,

$$\begin{aligned}\mu^k(Q^{(i,i')}) &\leq (1-\epsilon)\mu^k(R^{(i,i')})/(1-\epsilon/3) \\ &\leq (1-\epsilon/2)\mu^k(R^{(i,i')}).\end{aligned}$$

Let $Q = \bigcup_{(i,i') \in G} Q^{(i,i')}$ and $P = \bigcup_{(i,i') \notin G} R^{(i,i')}$. Then

$$\begin{aligned}\mu^k(Q) &\leq (1-\epsilon/2)\mu^k(\bigcup_{(i,i') \in G} R^{(i,i')}) \\ &= (1-\epsilon/2)\mu^k(A \times B - P - E).\end{aligned}$$

Furthermore for the projection P_1 of P on the first coordinate, $\mu(P_1) < K^2 2^{-m}$. Observe that the conditions that determine whether an element $(\vec{x}, \vec{y}) \in A \times B$ is in Q or P is based solely on the the (x_1, y_1) coordinates of $(\vec{x}, \vec{y})$ so each of Q and P is rectangular with respect to $T = \{2, \ldots, k\}$. □

Proof of Theorem 6. By symmetry we can assume without loss of generality that $b = 0$ and the first t_0 coordinates of v are 0. We will classify inputs in R based on the properties of their projections on each of the t_0 prefixes of their coordinates based on the trichotomy given by Lemma 7. Lemma 7 splits the set of inputs in any rectangle R based solely on their the first coordinate into a tiny error set E of inputs, a set P of inputs among which there are very few choices for the first coordinate and a set Q of the remaining inputs on which an output of 0 for that coordinate can be correct only on a $(1-\epsilon/2)$ fraction of inputs.

The sets of inputs corresponding to sets P and Q will be further subdivided using Lemma 7 based on the properties of their second coordinate, etc. For $j \leq t_0$ we will group together all the tiny error sets E found at any point into a single error set which also will be tiny. For the remaining inputs the decomposition over the various coordinates leads to disjoint sets of inputs corresponding to the branches of a binary tree, depending on whether the input fell into the P or Q set at each each application of Lemma 7. At each stage we either get a very small multiplicative factor in the upper bound on the total number of inputs possible because of the lack of variation in the coordinate (the case of set P) or we get a small multiplicative factor in the upper bound on the fraction of remaining inputs on which the answer of 0 can be correct (the case of set Q). For $\alpha \in \{p, q\}^{t_0}$ we will write S^α for the set of inputs such that for each $j \in [t_0]$, the input is in a P set at coordinate j when $\alpha_j = p$ and in a Q set at coordinate j when $\alpha_j = q$. Out of t_0 coordinates, one of p or q must occur at least $t_0/2$ times which will be good enough to derive the claimed bound.

Define μ^T for $T \subseteq [t]$ as in the proof of Lemma 7. For $\alpha \in \{p, q\}^j$ define $\#_p(\alpha)$ (resp. $\#_q(\alpha)$) to be the number

of p's (resp. q's) in α. For $0 \le j \le t_0$ and $\alpha \in \{p,q\}^j$ we will define sets $S^\alpha, E^j \subseteq X^t \times Y^t$, with the following properties for each $j \le t_0$:

1. $R \cap (f^t)^{-1}(v) \subseteq E^j \cup \bigcup_{\alpha \in \{p,q\}^j} S^\alpha$.
2. For every $\alpha \in \{p,q\}^j$, S^α is rectangular with respect to coordinates $j+1,\ldots,t$.
3. For $U = \{1,\ldots,j\}$, the sets S^α_U for different $\alpha \in \{p,q\}^j$ are disjoint.
4. For $\alpha \in \{p,q\}^{j-1}$, $\mu^t(S^{\alpha q}) \le (1-\epsilon/2)[\mu^t(S^\alpha) - \mu^t(S^{\alpha p})]$.
5. For $U = \{1,\ldots,j\}$, $\mu^U(S^\alpha_U) \le \lceil -mt/\log(1-\epsilon/6)\rceil^{2j} 2^{-\#_p(\alpha)m}$ for any $\alpha \in \{p,q\}^j$.
6. $\mu^t(E^j) \le 2j2^{-mt}$.

Define $S^\lambda = R$ and $E^0 = \emptyset$ where λ is the empty string. Clearly all the properties are satisfied for $j = 0$.

For each $\alpha \in \{p,q\}^j$ we apply Lemma 7 to build the sets $S^{\alpha p}$, $S^{\alpha q}$, and E^{j+1} from sets S^α and E^j as follows.

Let $\alpha \in \{p,q\}^j$. Let $U = \{1,\ldots,j\}$ and $T = [t] - U$. For each $(x_U, y_U) \in S^\alpha_U$, $S^\alpha(x_U,y_U)_T$ can be expressed as $A_{(x_U,y_U)} \times B_{(x_U,y_U)}$. Apply Lemma 7 with $k = t-j$ and $K' = mt$ to $A_{(x_U,y_U)} \times B_{(x_U,y_U)}$ to obtain disjoint sets $P_{(x_U,y_U)}$, $Q_{(x_U,y_U)}$, and $E_{(x_U,y_U)}$ that contain all inputs of $S^\alpha(x_U,y_U)$ on which the $j+1$-st output 0 is correct. (Lemma 7 yields sets P, Q, and E that are defined on coordinates $j+1,\ldots,t$ but we extend them to all coordinates by including (x_U,y_U) in all of the sets.) These sets are disjoint on coordinate j, rectangular on coordinates $j+2,\ldots,t$ and for $K = \lceil -mt/\log_2(1-\epsilon/6)\rceil$ satisfy

$$\begin{aligned}
&\mu^T((E_{(x_U,y_U)})_T) \le 2^{1-mt},\\
&\mu((P_{(x_U,y_U)})_{j+1}) \le K^2 2^{-m}, \text{ and}\\
&\mu^T((Q_{(x_U,y_U)})_T) \le (1-\epsilon/2)\\
&\qquad\qquad \times \mu^T(S^\alpha(x_U,y_U)_T - (P_{(x_U,y_U)})_T).
\end{aligned}$$

For $\alpha \in \{p,q\}^j$ define

$$S^{\alpha p} = \bigcup_{(x_U,y_U)\in S^\alpha_U} P_{(x_U,y_U)},$$
$$S^{\alpha q} = \bigcup_{(x_U,y_U)\in S^\alpha_U} Q_{(x_U,y_U)},$$

and define

$$E^{j+1} = E^j \cup \bigcup_{\alpha\in\{p,q\}^j} \bigcup_{(x_U,y_U)\in S^\alpha_U} E_{(x_U,y_U)}.$$

Properties 1, 2, and 3 for $j+1$ follow immediately from Lemma 7.

Furthermore,

$$\begin{aligned}
&\mu^t(S^{\alpha q})\\
&= \mu^t(\bigcup_{(x_U,y_U)\in S^\alpha_U} Q_{(x_U,y_U)})\\
&= \sum_{(x_U,y_U)\in S^\alpha_U} \mu^t(Q_{(x_U,y_U)})\\
&= \sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\})\mu^T(Q_{(x_U,y_U)})\\
&\le \sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\})\\
&\quad \times (1-\epsilon/2)\mu^T(S^\alpha(x_U,y_U)_T - (P_{(x_U,y_U)})_T)\\
&= (1-\epsilon/2)[\sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\})\mu^T(S^\alpha(x_U,y_U)_T\\
&\quad - \sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\})\mu^T(P_{(x_U,y_U)})_T)]\\
&= (1-\epsilon/2)[\mu^t(S^\alpha) - \mu^t(S^{\alpha p})]
\end{aligned}$$

which proves that property 4 is satisfied for $j+1$.

Also,

$$\begin{aligned}
&\mu^t(E^{j+1})\\
&= \mu^t(E^j \cup \bigcup_{\alpha\in\{p,q\}^j} \bigcup_{(x_U,y_U)\in S^\alpha_U} \mu^t(E_{(x_U,y_U)}))\\
&\le \mu^t(E^j) + \sum_{\alpha\in\{p,q\}^j} \sum_{(x_U,y_U)\in S^\alpha_U} \mu^t(E_{(x_U,y_U)})\\
&= \mu^t(E^j)\\
&\quad + \sum_{\alpha\in\{p,q\}^j} \sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\})\mu^T((E_{(x_U,y_U)})_T)\\
&\le 2j2^{-mt}\\
&\quad + \sum_{\alpha\in\{p,q\}^j} \sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\})\mu^T((E_{(x_U,y_U)})_T)\\
&\le 2j2^{-mt} + \mu^U(\bigcup_{\alpha\in\{p,q\}^j} S^\alpha_U)\, 2^{1-mt}\\
&\le 2j2^{-mt} + 2^{1-mt} \le 2(j+1)2^{-mt},
\end{aligned}$$

which proves that property 6 is satisfied for $j+1$.

Finally, for property 5 observe that for $\alpha \in \{p,q\}^j$,

$$\begin{aligned}
&\mu^{U\cup\{j+1\}}(S^{\alpha p}_{U\cup\{j+1\}})\\
&= \mu^{U\cup\{j+1\}}(\bigcup_{(x_U,y_U)\in S^\alpha_U} (P_{(x_U,y_U)})_{U\cup\{j+1\}})\\
&= \sum_{(x_U,y_U)\in S^\alpha_U} \mu^{U\cup\{j+1\}}((P_{(x_U,y_U)})_{U\cup\{j+1\}})\\
&= \sum_{(x_U,y_U)\in S^\alpha_U} \mu^U(\{(x_U,y_U)\}) \cdot \mu((P_{(x_U,y_U)})_{j+1})
\end{aligned}$$

$$
\begin{aligned}
&= \mu^U(S^\alpha_U) \cdot \mu((P_{(x_U,y_U)})_{j+1}) \\
&\le \mu^U(S^\alpha_U) \cdot K^2 2^{-m} \\
&\le K^{2j} 2^{-\#_p(\alpha)m} \cdot K^2 2^{-m} \\
&= K^{2(j+1)} 2^{-\#_p(\alpha p)m}
\end{aligned}
$$

and

$$
\begin{aligned}
\mu^{U\cup\{j+1\}}(S^{\alpha q}_{U\cup\{j+1\}}) &\le \mu^{U\cup\{j+1\}}(S^\alpha_{U\cup\{j+1\}}) \\
&= \mu^U(S^\alpha_U) \cdot \mu(S^\alpha_{j+1}) \\
&\le \mu^U(S^\alpha_U) \\
&\le K^{2j} 2^{-\#_p(\alpha)m} \\
&= K^{2j} 2^{-\#_p(\alpha q)m}.
\end{aligned}
$$

Thus property 5 is satisfied for $j+1$.

This implies all the properties required for the inductive hypothesis and we have produced the desired sets. We now use all these properties to derive the upper bound on $\mu^t(R \cap (f^t)^{-1}(v))$:

By property 1, $R \cap (f^t)^{-1}(v) \subseteq E^{t_0} \cup \bigcup_{\alpha\in\{p,q\}^{t_0}} S^\alpha$. Therefore for $\alpha \in \{p,q\}^{t_0}$ with $\#_p(\alpha) \ge t_0/2$,

$$
\begin{aligned}
\mu^t(S^\alpha) &\le \mu^{\{1,\ldots,t_0\}}(S^\alpha_{\{1,\ldots,t_0\}}) \\
&\le K^{2t_0} 2^{-\#_p(\alpha)m} \\
&\le K^{2t_0} 2^{-t_0 m/2}
\end{aligned}
$$

so $\mu^t(\bigcup_{\alpha\in\{p,q\}^{t_0}:\#_p(\alpha)\ge t_0/2} S^\alpha) \le 2^{t_0} K^{2t_0} 2^{-t_0 m/2}$.

We now upper bound the total measure of S^α for $\#_p(\alpha) \le t_0/2$.

CLAIM: For every $j \le t_0$, $\mu^t(\bigcup_{\alpha\in\{p,q\}^{t_0}:\,\#_q(\alpha)=j} S^\alpha) \le (1-\epsilon/2)^j \mu^t(R)$.

The claim is clearly true for $j=0$. For any $\alpha \in \{p,q\}^*$, by multiple applications of property 4,

$$
\begin{aligned}
&\mu^t(\bigcup_{i\le t_0-|\alpha|-1} S^{\alpha p^i q}) \\
&= \sum_{i\le t_0-|\alpha|-1} \mu^t(S^{\alpha p^i q}) \\
&\le \sum_{i\le t_0-|\alpha|-1} (1-\epsilon/2)[\mu^t(S^{\alpha p^i}) - \mu^t(S^{\alpha p^{i+1}})] \\
&\le (1-\epsilon/2)\mu^t(S^\alpha)
\end{aligned}
$$

since the sum telescopes. Let $A_j = (p^*q)^j \cap \{p,q\}^{\le t_0}$ be the set of all strings of length up to t_0 that end in a q and have a total of j q's. The above for $\alpha = \lambda$ implies that $\mu^t(\bigcup_{\alpha'\in A_1} S^{\alpha'}) \le (1-\epsilon/2)\mu^t(R)$. We can also apply the above to all $\alpha \in A_j$ to yield that $\mu^t(\bigcup_{\alpha'\in A_{j+1}} S^{\alpha'}) \le (1-\epsilon/2)\mu^t(\bigcup_{\alpha\in A_j} S^\alpha)$ and thus by induction that $\mu^t(\bigcup_{\alpha\in A_j} S^\alpha) \le (1-\epsilon/2)^j \mu^t(R)$. Finally, since $S^{\alpha p} \subseteq S^\alpha$ for any α we derive that

$$
\begin{aligned}
\mu^t(\bigcup_{\alpha\in\{p,q\}^{t_0}:\,\#_q(\alpha)=j} S^\alpha) &= \mu^t(\bigcup_{\alpha\in A_j} S^{\alpha p^{t_0-|\alpha|}}) \\
&\le \mu^t(\bigcup_{\alpha\in A_j} S^\alpha) \\
&\le (1-\epsilon/2)^j \mu^t(R)
\end{aligned}
$$

and the claim is proved.
Thus the total

$$
\begin{aligned}
&\mu^t(\bigcup_{\alpha\in\{p,q\}^{t_0}:\,\#_p(\alpha)<t_0/2} S^\alpha) \\
&= \mu^t(\bigcup_{\alpha\in\{p,q\}^{t_0}:\,\#_q(\alpha)>t_0/2} S^\alpha) \\
&\le (2/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R).
\end{aligned}
$$

Putting it all together we have

$$
\begin{aligned}
&\mu^t(R \cap (f^t)^{-1}(v)) \\
&\le \mu^t(E^{t_0}) + \mu^t(\bigcup_{\alpha\in\{p,q\}^{t_0}} S^\alpha) \\
&= \mu^t(E^{t_0}) + \mu^t(\bigcup_{\substack{\alpha\in\{p,q\}^{t_0} \\ \#_p(\alpha)\ge t_0/2}} S^\alpha) + \mu^t(\bigcup_{\substack{\alpha\in\{p,q\}^{t_0} \\ \#_p(\alpha)<t_0/2}} S^\alpha) \\
&\le 2t_0 2^{-mt} + 2^{t_0} K^{2t_0} 2^{-t_0 m/2} + (2/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R).
\end{aligned}
$$

Since $-\log_2(1-\epsilon/6) > -\sqrt{2}\ln(1-\epsilon/6) \ge \sqrt{2}\epsilon/6$ and $\epsilon > 6mt/2^{m/8}$, $K = \lceil -mt/\log_2(1-\epsilon/6)\rceil < 2^{m/8}/2^{3/2}$ and therefore

$$
2^{t_0} K^{2t_0} 2^{-t_0 m/2} < 2^{-t_0 m/4}/2^{2t_0}.
$$

Therefore if $\mu^t(R) \ge 2^{-t_0 m/6}$ and, since the condition on ϵ implies that $m \ge 24$,

$$
\begin{aligned}
&\mu^t(R \cap (f^t)^{-1}(v)) \\
&< 2t_0 2^{-mt} + 2^{-t_0 m/4}/2^{2t_0} + (2/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R) \\
&< 2 - t_0 m/4 + (2/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R) \\
&\le 2^{-t_0 m/24}\mu^t(R) + (2/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R) \\
&\le 2^{-t_0}\mu^t(R) + (2/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R) \\
&\le (3/\epsilon)(1-\epsilon/2)^{t_0/2}\mu^t(R)
\end{aligned}
$$

as required. □

The following is a direct sum theorem for randomized communication complexity derived from corruption bounds on cross product distributions on rectangles.

Theorem 8. *Let $f : X \times Y \to \{0,1\}$ and let μ be a rectangular distribution on $X \times Y$. For $b \in \{0,1\}$, $p = \mu(f^{-1}(b))$, $\epsilon < p$, if t is an integer such that $\log_2 t \leq \mathrm{corrbd}^b_\mu(f,\epsilon)/32$ and $\epsilon \geq 9\ln(pt)/pt$ then there are constants $c, c', c'' > 0$ such that*

$$R^2_{1-(1-\epsilon)^{c''pt}}(f^t) \geq cpt \cdot \mathrm{corrbd}^b_\mu(f,\epsilon) - c'pt.$$

Proof. Without loss of generality for ease of notation we assume $b = 0$, write m for $\mathrm{corrbd}^0_\mu(f,\epsilon)$ and assume that m is sufficiently large.

We will apply Lemma 4 with the set $O' = \{v \in \{0,1\}^t \mid v \text{ has} \geq pt/4\ 0\text{'s}\}$. Let I_s be the set of all inputs $(\vec{x},\vec{y}) \in X^t \times Y^t$ such that $f^t(\vec{x},\vec{y})$ contains precisely s 0's. By definition $\mu^t(I_s) = \Pr[B(t,p) = s]$ where $B(t,p)$ is the binomial distribution that is the sum of t Bernoulli trials with success probability p. Therefore $\mu^t(\bigcup_{s<pt/4} I_s) \leq 2^{-pt/2}$ and thus $\mu(f^{-1}(O')) \geq 1 - 2^{-pt/2}$.

For $\log_2 t \leq m/32$, $\epsilon \geq 9\ln(pt)/pt > 1/t$, for sufficiently large m, $\epsilon \geq 12mt/2^{m/8}$ and $(4/\epsilon)(1-\epsilon/2)^{pt/4} < (1-\epsilon)^{c''pt}$ for some constant $c'' > 0$. We can then apply Theorem 6 to show that for every $v \in O'$, $\mathrm{corrbd}^v_{\mu^t}(f^t, 1 - (3/\epsilon)(1-\epsilon/2)^{pt/4}) \geq ptm/24$. By Lemma 4, for $\delta = (3/\epsilon)(1-\epsilon/2)^{pt/4}$ and $\delta' < 1$,

$$R^2_{1-\delta'}(f) \geq \frac{ptm}{24} - \log_2\left(\frac{1}{\mu(f^{-1}(O') - (1-\delta')/(1-\delta)}\right).$$

Since $\epsilon < p$, for $\delta' = (4/\epsilon)(1-\epsilon/2)^{pt/4}$, $\delta' \geq \delta + 2^{1-pt/2}$ and $(1-\delta')/(1-\delta) \leq 1 - 2^{1-pt/2}$. Thus $\mu(f^{-1}(O') - (1-\delta')/(1-\delta) \geq 2^{-pt/2}$ by our lower bound on $\mu(f^{-1}(O'))$. Therefore $R^2_{1-\delta'}(f) \geq ptm/24 - pt/2$. We choose $c = 1/24$ and $c' \geq 1/2$ to ensure that the bound is non-trivial only when m is large enough for the above conditions on m to hold. This yields the claimed bound. □

Set Disjointness

Define the k-party set-disjointness function for $X_1 = \cdots = X_k = \{0,1\}^n$ by $\mathrm{DISJ}_{k,n} : X_1 \times \cdots \times X_k \to \{0,1\}$ by $\mathrm{DISJ}_{k,n}(x_1,\ldots,x_k) = 1$ if for there is some $j \in [n]$ such that $x_{i,j} = 1$ for all $i \in [k]$ and $\mathrm{DISJ}_{k,n}(x_1,\ldots,x_k) = 0$ otherwise. We drop the subscript n if it is understood from the context.

Babai, Frankl, and Simon [2] obtained a 2-party randomized communication complexity lower bound for DISJ_2 by showing that for some constant $\epsilon > 0$ $\mathrm{corrbd}^0_\mu(\mathrm{DISJ}_2,\epsilon)$ is $\Omega(\sqrt{n})$ and $p = \mu(f^{-1}(0))$ is constant where μ is the rectangular distribution given by $\Pr_\mu[x_i = 1] = \Pr_\mu[y_i = 1] = n^{-1/2}$ independently. Theorem 8 immediately allows us to derive the following bound from this result:

Corollary 9. *For some constant $\delta < 1$, for any integer t, $R^2_{1-\delta^t}(\mathrm{DISJ}^t_2)$ is $\Omega(t\sqrt{n})$.*

Remark 1. *Using the direct sum property for conditional information complexity and the lower bound of [7], for fixed error $\epsilon < 1$ one can obtain the bound $R^2_\epsilon(\mathrm{DISJ}^t_2)$ is $\Omega(tn)$. However this bound is incomparable to the above corollary which allows exponentially small correctness.*

In general we are interested in the complexity of $R^k_\epsilon(\mathrm{DISJ}_k)$. A natural distribution on $X_1 \times \cdots \times X_k$ to consider is ν_k under which $\Pr_{\nu_k}[x_{i,j} = 1] = n^{-1/k}$ for every i and j independently. This is the natural extension to DISJ_k of the distribution of [2],

4 3-party NOF Communication Complexity of Disjointness

Here we consider the computation of DISJ_3 and write X, Y, Z for the 3 components of the domain. In the NOF (number-on-the-forehead) model we identify 3 parties, the X player who receives (y,z) as input, the Y player who receives (x,z) as input and the Z player who receives (x,y) as input. The natural distribution $\nu = \nu_3$ on $X \times Y \times Z$ has, independently, $\Pr_\nu[x_i = 1] = \Pr_\nu[y_i = 1] = \Pr_\nu[z_i = 1] = n^{-1/3}$.

4.1 $Z \to (Y \leftrightarrow X)$ Computation

As an approach for obtaining size-depth tradeoff lower bounds in circuit complexity Nisan and Wigderson [24] suggested analyzing the 3-party NOF complexity of functions of the form $f(x,h,i) = h(x)_i$ for H a family of universal hash functions $h : X \to \{0,1\}^n$. They showed that sufficiently strong lower bounds on the communication complexity when the i-player holding (h,x) communicates first (and then ceases interaction) would suffice to derive functions not computable in $O(\log n)$ depth and linear size simultaneously.

We analyze the (randomized) communication complexity of set disjointness in this model which we write mnemonically as $D^{Z\to(Y\leftrightarrow X)}(\mathrm{DISJ}_3)$ and $R^{Z\to(Y\leftrightarrow X)}_\epsilon(\mathrm{DISJ}_3)$.

Theorem 10. *$D^{Z\to(Y\leftrightarrow X)}(\mathrm{DISJ}_3)$ is $\Omega(n^{1/3})$ and for $\epsilon < 1/2$, $R^{Z\to(Y\leftrightarrow X)}_\epsilon(\mathrm{DISJ}_3)$ is $\Omega((1-2\epsilon)n^{1/3}/\log n)$.*

Proof. We follow the general approach of [24] but use a direct sum bound for corruption in place of a discrepancy bound for universal hash function families. Note that although the basic approach and bound of [24] is correct, there is an issue with the proof in [24] that is discussed and corrected below.

Fix any $Z \to (Y \leftrightarrow X)$ protocol P computing DISJ_3 and let $C(P)$ be the total number of bits communicated in P. Let $t = n^{1/3}$. View each string x, y, z as a sequence of t blocks, $x_1,\ldots,x_t, y_1,\ldots,y_t, z_1,\ldots,z_t \in \{0,1\}^{n/t}$.

Given P we first construct a $Z \to (Y \leftrightarrow X)$ protocol P' that computes $(\textsc{Disj}_2(x_1, y_1), \ldots, \textsc{Disj}_2(x_t, y_t))$ in which the Z-player sends $C(P)$ bits and the X and Y players together send $tC(P)$ bits: Consider runs of the protocol P with different choices of $z \in Z$, in particular with $z = z^j = 0^{(j-1)n/t}1^{n/t}0^{(t-j)n/t}$ for $j = 1, \ldots, t$. For $z = z^j$, $\textsc{Disj}_3(x, y, z) = \textsc{Disj}_2(x_j, y_j)$. Also observe that for each of these choices, the message $m_Z(x, y)$ sent by the Z-player is independent of the choice of z. On input (x, y), the new protocol P' simulates P on inputs (x, y, z^j) for $j = 1, \ldots, t$ except that, since the message sent by the Z-player is the same in each case, the Z-player sends this message only once. P' then outputs the tuple of results.

Observe that the function computed by P' does not depend on the choice of z so we define a new protocol $P''(x, y) = P'(x, y, 0^n)$. In protocol P'' the Z-player receives (x, y) as input as before but the X player only receives x and the Y player only receives y. (We swap the names of the X and Y players to avoid confusion.) After the Z-player's communication of $C(P)$ bits, the X- and Y-players exchange $tC(P)$ bits in order to compute $\textsc{Disj}^t_{2,n/t}(x, y)$.

Sequentially fix the communications of the players as follows: Consider the distribution ν on $X \times Y \times Z$ defined above and let μ be the induced distribution on x and y. Let $p = \Pr_\mu[\textsc{Disj}_2(x^j, y^j) = 0]$ be the probability that x and y intersect in block j. By construction, p is an absolute constant. For a vector v, write $\#_0(v)$ to be the number of 0's in v. Since P'' is always correct, by Chernoff bounds

$$\Pr_\mu[\#_0(P''(x, y)) < \frac{pt}{4}] = \Pr_\mu[\#_0(\textsc{Disj}^t_{2,n/t}(x, y)) < \frac{pt}{4}] < 2^{-pt/2} \leq 1/2.$$

Since the set of possible messages is prefix-free and $|m_z| \leq C(P)$, there is some m_z such that

$$\Pr_\mu[m_Z(x, y) = m_z \text{ and } \#_0(P''(x, y)) \geq \frac{pt}{4}] > 2^{-C(P)-1}.$$

Fix that m_z.

At this point in [24], the communications of the X and Y players are fixed to frequent strings m_x and m_y and the claim is made that the set of inputs on which the communications m_x and m_y occur is a rectangle. Unfortunately, this is not necessarily the case; for example, it is violated by a protocol in which $m_z = 1$ if and only if $x_1 = y_1$ and the X-player and Y-player toggle each bit they communicate depending on m_z. Instead we apply a more complicated argument.

Let $S_{m_z} \subseteq X \times Y$ be the set of inputs on which $m_Z(x, y) = m_z$. For a possible communication sequence m_{XY} of the X and Y players we say that x is *consistent with* (m_z, m_{XY}) if and only if there is some y such that $(x, y) \in S_{m_z}$ and the communication in P'' by the X- and Y-players is m_{XY} after the Z-player has sent m_z. The definition is analogous for y. Let
$A_{m_{XY}}(m_z) = \{x \in X \mid x \text{ is consistent with } (m_z, m_{XY})\}$,
$B_{m_{XY}}(m_z) = \{y \in Y \mid y \text{ is consistent with } (m_z, m_{XY})\}$.

By the usual argument for 2-party protocols, for m_z fixed, the different rectangles $R_{m_{XY}}(m_z) = A_{m_{XY}}(m_z) \times B_{m_{XY}}(m_z)$ are disjoint; moreover the communication under P'' is (m_z, m_{XY}) for every input in $R_{m_{XY}}(m_z) \cap S_{m_z}$. For simplicity write $R_{m_{XY}}$ for $R_{m_{XY}}(m_z)$. Let $S \subseteq S_{m_z}$ be the portion of S_{m_z} on which $\#_0(P''(x, y)) \geq pt/4$. By construction $\mu(S) \geq 2^{-C(P)-1}$.

There are at most $2^{tC(P)}$ valid choices for m_{XY}. By Markov's inequality, at least $1/2$ of the measure of S has communication $m_{XY} = m_{xy}$ such that

$$\mu(R_{m_{xy}} \cap S) \geq 2^{-tC(P)-1}\mu(S) \geq 2^{-(t+1)C(P)-2}.$$

For any such m_{xy} clearly $\mu(R_{m_{xy}}) \geq 2^{-(t+1)C(P)-2}$. For such an m_{xy}, let the output of P'' with communication given by (m_z, m_{xy}) be v. By construction, $\#_0(v) \geq pt/4$.

By Theorem 6 applied to $\textsc{Disj}_{2,n/t}$ there are constants $\delta < 1$ and $c' > 0$ such that $\mathrm{corrbd}^v_\mu(\textsc{Disj}^t_{2,n/t}, 1 - \delta^t) \geq c't\sqrt{n/t}$. Therefore for $(t + 1)C(P) + 2 \leq c't\sqrt{n/t}$, by definition,

$$\Pr_\mu[\textsc{Disj}^t_{2,n/t}(x, y) = v \mid (x, y) \in R_{m_{xy}}] \leq \delta^t.$$

By the assumption that P'' is always correct, $\textsc{Disj}^t_{2,n/t}(x, y) = v$ for all $(x, y) \in R_{m_{xy}} \cap S$. Thus $\mu(R_{m_{xy}} \cap S) \leq \delta^t\mu(R_{m_{xy}})$. Therefore if $(t+1)C(P)+2 \leq c't\sqrt{n/t}$ then $2^{-C(P)-2} \leq \mu(S)/2 \leq \delta^t$ since the various rectangles $R_{m_{xy}}$ are disjoint. It follows that $C(P)$ is $\Omega(\min\{t, \sqrt{n/t}\})$ which is $\Omega(n^{1/3})$ since $t = n^{1/3}$.

The case of randomized complexity is very similar although a little more complicated. We first repeat the protocol P in parallel and take the majority answer to reduce its error from ϵ to $1/(8t)$. This increases the communication complexity $C(P)$ by a factor that is $O(\frac{\log t}{1-2\epsilon})$. We then use Yao's lemma with the distribution ν to derive a deterministic protocol P^* with complexity $C(P^*)$ that is $O(\frac{\log t}{1-2\epsilon}C(P))$ and has error at most $1/(8t)$ on the distribution ν.

We apply the above argument with P^* replacing P to obtain a protocol P'' computing $\textsc{Disj}^t_{2,n/t}(x, y)$ in which the Z-player sends $C(P^*) = O(\frac{\log t}{1-2\epsilon}C(P))$ bits based on (x, y) and the X and Y players interact sending a total of $tC(P^*)$ bits based on x and y respectively. The error of P'' is the probability under ν (and therefore under μ since P'' does not depend on the value of z) that any one of the

t components of its answer is incorrect which is at most $t \cdot 1/(8t) = 1/8$ by a union bound. In this case,

$$\begin{aligned}
&\Pr_{\mu}[\#_0(P''(x,y)) \geq pt/4 \text{ and } P''(x,y) = \textsc{Disj}^t_{2,n/t}(x,y)] \\
&\geq \Pr_{\mu}[P''(x,y) = \textsc{Disj}^t_{2,n/t}(x,y)] \\
&\quad - \Pr_{\mu}[\#_0(\textsc{Disj}^t_{2,n/t}(x,y) < pt/4)] \\
&> 7/8 - 2^{-pt/2} > 3/4.
\end{aligned}$$

Each possible communication string $m = (m_Z, m_{XY})$ is associated with a set of inputs $T_m \subseteq X \times Y$ on which the output of P'' is constant $v(m)$ and these sets partition $X \times Y$. Since at most $1/4$ of the inputs in $X \times Y$ have either an error in P'' or few 0's in the output of P'', by Markov's inequality at least $1/2$ the measure of inputs in $X \times Y$ are contained in sets T_m for which the output of P'' on the inputs in T_m has at least $pt/4$ zeros and the probability that this output is correct is at least $1/2$ the measure of T_m.

Therefore since there are only $2^{C(P^*)}$ choices of m_Z, we can fix an m_z such that the total μ measure of sets T_m on which $m_Z(x,y) = m_z$, $\#_0(v(m)) \geq pt/4$, and $\Pr_\mu[P''(x,y) = \textsc{Disj}^t_{2,n/t}(x,y) \mid (x,y) \in T_m] \geq 1/2$ is at least $2^{-C(P^*)-1}$. Let S_{m_z} be the set of inputs consistent with communication m_z and let $S \subseteq S_{m_z}$ be the set of inputs (x,y) such that for some m_{XY} and $m = (m_z, m_{XY})$, P'' has communication m on input (x,y), $\#_0(v(m)) \geq pt/4$, and $\Pr_\mu[P''(x,y) = \textsc{Disj}^t_{2,n/t}(x,y) \mid (x,y) \in T_m] \geq 1/2$. Clearly $\mu(S) \geq 2^{-C(P^*)-1}$.

Again there are at most $2^{tC(P^*)}$ possible communication strings m_{XY} between the X and Y players in P'' that, together with m_z, define the set T_m. As above, each choice of m_{XY} yields a rectangle $R_{m_{XY}}(m_z)$ such that $T_m = R_{m_{XY}}(m_z) \cap S$. By the same reasoning as above at least a $1/2$ fraction of S is covered by disjoint rectangles $R_{m_{xy}}(m_z)$ such that $\mu(R_{m_{xy}}(m_z) \cap S) \geq \mu(S)2^{-tC(P^*)-1} \geq 2^{-(t+1)C(P^*)-2}$ and thus $\mu(R_{m_{xy}}(m_z)) \geq 2^{(t+1)C(P^*)-2}$. For any such choice of $m_{XY} = m_{xy}$ and $m = (m_z, m_{xy})$, and the correctness of P'' on T_m implies that

$$\begin{aligned}
&\Pr_{\mu}[\textsc{Disj}^t_{2,n/t}(x,y) = v \text{ and } (x,y) \in T_m] \\
&\geq \mu(T_m)/2 \;=\; \mu(R_{m_{xy}} \cap S)/2.
\end{aligned}$$

However,

$$\begin{aligned}
&\Pr_{\mu}[\textsc{Disj}^t_{2,n/t}(x,y) = v(m) \text{ and } (x,y) \in T_m] \\
&\leq \Pr_{\mu}[\textsc{Disj}^t_{2,n/t}(x,y) = v(m) \text{ and } (x,y) \in R_{m_{xy}}(m_z)] \\
&\leq \delta^t \mu(R_{m_{xy}}(m_z))
\end{aligned}$$

for $(t+1)C(P) + 2 \leq c't\sqrt{n/t}$ and some $\delta < 1$ by Theorem 6 applied to $\textsc{Disj}^t_{2,n/t}$. Since the various $R_{m_{xy}}$ are disjoint and $\mu(\bigcup_{m_{xy}}(R_{m_{xy}}(m_z) \cap S)) \geq \mu(S)/2 \geq 2^{-C(P^*)-2}$, summing up over all m we obtain that if $(t+1)C(P) + 2 \leq c't\sqrt{n/t}$ then

$$\begin{aligned}
2^{-C(P^*)-2} &\leq \mu(S)/2 \\
&\leq \mu(\bigcup_{m_{xy}}(R_{m_{xy}}(m_z) \cap S)) \\
&\leq \delta^t \mu(\bigcup_{m_{xy}} R_{m_{xy}}(m_z)) \;\leq \delta^t,
\end{aligned}$$

we obtain that $C(P^*)$ is $\Omega(t)$.

Thus $C(P^*)$ is $\Omega(\min\{t, \sqrt{n/t}\})$. This implies that $\frac{\log t}{1-2\epsilon}C(P)$ is $\Omega(\min\{t, \sqrt{n/t}\})$. Since $t = n^{1/3}$ this yields the claimed bound. □

4.2 General 3-party NOF Computation

In this we prove an $\Omega(\log n)$ lower bound on the unrestricted 3-party NOF computation of set disjointness. Although this is not yet strong enough to imply lower bounds for lift-and-project proof systems it is of independent interest since it is the first lower bound for general multiparty NOF communication complexity proved via a corruption bound that does not follow from a discrepancy bound.

Theorem 11. *For any $\epsilon' < 1/2$, $R^3_{\epsilon'}(\textsc{Disj}_3)$ is $\Omega((1-2\epsilon')^2 \log n)$.*

To prove this theorem we use the following simple characterization of 3-cylinder intersections.

Proposition 12. *A set E is a 3-cylinder intersection on $X \times Y \times Z$ if and only if there is a set $S \subseteq X \times Y$ and for each $z \in Z$ there is a combinatorial rectangle $R_z \in \mathcal{P}(X) \times \mathcal{P}(Y)$ such that $E = \bigcup_{z \in Z}((R_z \cap S) \times \{z\})$.*

Proof. By definition, a 3-cylinder intersection E consists of the intersection of an X-cylinder, a Y-cylinder, and a Z-cylinder. We can write the Z-cylinder as $S \times Z$ for some $S \subseteq X \times Y$. Under permutation of the components we can write the X- and Y-cylinders as $T \times X$ and $U \times Y$ respectively for $T \subseteq Y \times Z$ and $U \subseteq X \times Z$. For each $z \in Z$ write $T_z = \{y \in Y \mid (y,z) \in T\}$ and $U_z = \{x \in X \mid (x,z) \in U\}$. Define $R_z = U_z \times T_z$. Clearly R_z is a rectangle on $X \times Y$. Moreover by definition, for each $z \in Z$, $(x,y,z) \in E$ if and only if $(x,y) \in S$, $y \in T_z$ and $x \in U_z$; i.e., for each $z \in Z$, $(x,y,z) \in E$ if and only if $(x,y) \in S$ and $(x,y) \in R_z$, as required. □

Proof of Theorem 11. Let $t = n^{1/3}$. Define a distribution ν on $X \times Y \times Z$ as follows: Choose z uniformly at random from $\{z^j = 0^{(j-1)(n/t)}1^{n/t}0^{(t-j)n/t} \mid j \in [t]\}$ and independently choose each bit of x and y independently as

above, with $\Pr_\nu[x_i = 1] = \Pr_\nu[y_i = 1] = n^{-1/3}$. Clearly $\nu(\textsc{Disj}_3^{-1}(0)) = p$ where p is an absolute constant as above.

Let Γ be the set of all cylinder intersections on $X \times Y \times Z$. We prove that for some $\gamma > 0$ and any $\epsilon < 1$, $\epsilon\text{-mono}^0_{\nu,\Gamma} \leq n^{-\gamma}$. The claimed lower bound then follows by repeating the protocol $O(\frac{1}{(1-2\epsilon')^2})$ times to reduce the error below $p/2$ and then applying Corollary 5.

Consider any cylinder intersection E on $X \times Y \times Z$ and write $E = \bigcup_{z \in Z}((S \cap R_z) \times \{z\})$ for $S \subseteq X \times Y$ and R_z rectangles on $X \times Y$. Suppose that $\nu(\textsc{Disj}_3^{-1}(1) \cap E) \leq \epsilon \cdot \nu(E)$. It is sufficient prove that $\nu(E) \leq n^{-\gamma}$.

Because of the definition of ν we can assume without loss of generality that $E = \bigcup_{j=1}^t (S \cap R_{z^j}) \times \{z^j\}$. Let μ be the measure induced on $X \times Y$ by ν. For each $(x, y) \in S$ let $J_{(x,y)} \subseteq [t]$ be the set of $j \in [t]$ for which $(x, y) \in R_{z_j}$ and $\textsc{Disj}_3(x, y, z^j) = 0$. This implies that $\textsc{Disj}_2(x_j, y_j) = 0$ for all $j \in J(x, y)$. Let $t_0 = \lceil \frac{(1-\epsilon)\nu(E)t}{2} \rceil$ and let

$$S' = \{(x, y) \in S \mid |J_{(x,y)}| \geq t_0\}.$$

By the error assumption for E, $\nu(E \cap \textsc{Disj}_3^{-1}(0)) \geq (1 - \epsilon)\nu(E)$. Let $E' = \{(x, y, z^j) \in E \mid (x, y) \in S'\}$ be the of elements of E whose (x, y) components are in S'. By definition,

$$\begin{aligned}\nu((E - E') \cap \textsc{Disj}_3^{-1}(0)) &\leq \frac{t_0 - 1}{t}\mu(S - S') \\ &< \frac{(1-\epsilon)\nu(E)t/2}{t}\mu(S) \\ &\leq (1 - \epsilon)\nu(E)/2.\end{aligned}$$

Therefore $\nu(E') \geq \nu(E' \cap \textsc{Disj}_3^{-1}(0)) \geq (1 - \epsilon)\nu(E)/2$ and thus $\mu(S') \geq \nu(E') \geq (1 - \epsilon)\nu(E)/2$.

We now consider a rectangular refinement of the rectangles R_{z_j}. For $j = 1, \ldots, t$ write $R_{z_j} = A_j \times B_j$ for $A_j \subseteq X$ and $B_j \subseteq Y$. For $\alpha, \beta \in \{0, 1\}^t$ define the rectangle

$$R_{\alpha,\beta} = (\bigcap_{j:\alpha_j=1} A_j \cap \bigcap_{j:\alpha_j=0} \overline{A_j}) \times (\bigcap_{j:\beta_j=1} B_j \cap \bigcap_{j:\beta_j=0} \overline{B_j}).$$

By definition, these 2^{2t} rectangles are mutually disjoint. Every $(x, y) \in S$ is in some $R_{\alpha,\beta}$ for which $\alpha_j = \beta_j = 1$ for all $j \in J_{(x,y)}$ and thus $R_{\alpha,\beta} \subseteq \bigcap_{j \in J_{(x,y)}} R_{z_j}$. In particular, by definition $\textsc{Disj}_3(x, y, z^j) = 0$ for all $j \in J_{(x,y)}$ and thus $\textsc{Disj}_3(x, y, z^j) = 0$ for all $j \in J_{(x,y)}$. Therefore each $R_{\alpha,\beta}$ that contains a point of S' has an associated set of t_0 values $j \in [t]$ such that $\textsc{Disj}_3(x, y, z^j) = 0$ for all elements of $R_{\alpha,\beta} \cap S'$. This implies that there is a fixed set of at least t_0 outputs of $\textsc{Disj}^t_{2,n/t}(x, y)$ that are 0 for all elements of $R_{\alpha,\beta} \cap S'$.

By Theorem 6 and the corruption bound for 2-party disjointness, there are some constants $c, \delta > 0$ and such that for any α, β if $\mu(R_{\alpha,\beta}) \geq 2^{-ct_0\sqrt{n/t}}$ then $\mu(R_{\alpha,\beta} \cap S') \leq \delta^{t_0}\mu(R_{\alpha,\beta})$. At most $2^{2t-ct_0\sqrt{n/t}}$ measure of points in S' can be covered by rectangles $R_{\alpha,\beta}$ for which $\mu(R_{\alpha,\beta}) < 2^{-ct_0\sqrt{n/t}}$. Since the rectangles $R_{\alpha,\beta}$ covering S' are disjoint, by the corruption bound the total measure of the part of S' covered by rectangles $R_{\alpha,\beta}$ with $\mu(R_{\alpha,\beta}) \geq 2^{-ct_0\sqrt{n/t}}$ is at most δ^{t_0}. Therefore $\mu(S') \leq \delta^{t_0} + 2^{2t-ct_0\sqrt{n/t}}$ which, for $t = n^{1/3}$, is at most $\delta^{t_0} + 2^{-(ct_0-2)t}$. Therefore $(1 - \epsilon)\nu(E)/2 \leq \delta^{t_0} + 2^{-(ct_0-2)t}$. Since $t_0 \geq (1 - \epsilon)\nu(E)t/2$, $\nu(E)$ is $O(\frac{\log t}{t})$ which is $O(\frac{\log n}{n^{1/3}})$ as required. □

Observe that the corruption bound under the distribution used in the proof of Theorem 11 is asymptotically tight: The X or Y player simply has to send $\lceil \log_2 t \rceil$ bits specifying the value of j and then the Z player can simply compute $\textsc{Disj}_3(x, y, z^j)$.

There are natural distributions one could consider, such as the distribution ν_3. The proof is a little more involved but it is possible to obtain similar corruption bounds for cylinder intersections under ν_3. These alternate distributions have potential utility in deriving much larger lower bounds. For example, we would like to improve the lower bound to something that is $\Omega(n^\delta)$ for some $\delta > 0$. The key limitation of the method of proof of Theorem 11 is the step in which we create the rectangular refinement of the set of rectangles.

5 k-party NOF Communication Complexity of Disjointness

5.1 Simultaneous k-party NOF Computation

The communication complexity in the NOF simultaneous messages case can be analyzed using the techniques of Babai, Gal, Kimmel and Lokam [3]. In this model the inputs are distributed in the usual NOF fashion but there is no interaction. Each player sends a single message to a referee who uses their contents to evaluate the function. We write $D^{X_1||\cdots||X_k}(f)$ and $R^{X_1||\cdots||X_k}_\epsilon(f)$ for the deterministic and ϵ-error randomized simultaneous NOF communication complexity of f.

Following [3] we directly analyze the complexity of this problem when one of the players, say player k, acts as the referee. We denote such protocols as $(X_1|| \cdots ||X_{k-1}) \to X_k$ protocols and, for example, use $R^{(X_1||\cdots||X_{k-1})\to X_k}_\epsilon(f)$ to denote the complexity in this case. Clearly, $R^{X_1||\cdots||X_k}_\epsilon(f) \geq R^{(X_1||\cdots||X_{k-1})\to X_k}_\epsilon(f)$.

The general idea of the approach in [3] is to find a small collection of possible inputs Q_i in each of the components X_i for $i \in [k - 1]$ with the property that taking all their

combinations together yields a large number of different subproblems player k might need to solve. The only information that player k receives about x_k is from the other players so the information from all their possible messages must be enough to differentiate among these possibilities.

Let $X_1 = \ldots = X_k = \{0,1\}^n$ which we also identify with $\mathcal{P}(n)$ and we write $x \cap y$ for the string whose i-th coordinate is 1 if and only if the i-th coordinate of both x and y are 1. Thus $\textsc{Disj}_k(x_1, \ldots, x_k) = 1$ if and only if $x_1 \cap \cdots \cap x_k \neq \emptyset$. We will refer to elements of $\{0,1\}^n$ interchangeably as sets or vectors. For C and D subsets of $\{0,1\}^n$ write $C \sqcap D = \{x \cap y \mid x \in C,\ y \in D\}$.

Proposition 13. *For $\ell \geq 1$ there exist $Q_1, \ldots, Q_\ell \subseteq \{0,1\}^n$ such that $|Q_i| = n^{1/\ell}$ and $Q_1 \sqcap \cdots \sqcap Q_\ell$ is the set of all singleton sets in $[n]$.*

Proof. Let $m = n^{1/\ell}$ and view $[n]$ as an ℓ-dimensional cube of side m. Let $Q_i = \{Q_{i1}, \ldots, Q_{im}\}$ be the partition of $[n]$ into subsets of size $m^{\ell-1}$ given by the m layers along the i-th dimension in this cube. Since the different sets within each Q_i are disjoint, all-nonempty sets in $Q_1 \sqcap \cdots \sqcap Q_\ell$ are disjoint. An element $j \in [n]$ can be indexed by its coordinates $(j_1, \ldots, j_\ell)$ in each of the ℓ dimensions of this cube. Clearly $\{j\} = Q_{1j_1} \cap Q_{2j_2} \cap \cdots \cap Q_{\ell j_\ell}$. □

For $0 \leq \epsilon \leq 1$ define $H_2(\epsilon) = \epsilon \log_2 \frac{1}{\epsilon} + (1-\epsilon) \log_2 \frac{1}{1-\epsilon}$ and let H be the binary entropy function.

Theorem 14.

$$R_\epsilon^{(X_1||\cdots||X_{k-1})\to X_k}(f) \geq (1 - H_2(\epsilon)) n^{1/(k-1)}/(k-1).$$

Proof. We apply Yao's lemma and analyze the complexity $C(P)$ of an ϵ-error deterministic protocol P under distribution μ given as follows: Apply Proposition 13 with $\ell = k-1$ to obtain sets $Q_1, \ldots, Q_{k-1} \subseteq \{0,1\}^n$ with $|Q_i| = m = n^{1/(k-1)}$ such that $Q_1 \sqcap \cdots \sqcap Q_{k-1}$ contains all singleton subsets of $[n]$. For each $j \in [n]$ we can identify a (unique) tuple $\vec{x}^j = (x_1^j, \ldots, x_{k-1}^j) \in Q_1 \times \cdots \times Q_{k-1}$ such that $\{j\} = x_1^j \cap \cdots \cap x_{k-1}^j$. Define distribution μ on $X_1 \times \ldots \times X_k$ by by choosing j uniformly at random from $[n]$ and independently choosing a uniformly random subset $x_k \subseteq [n]$ to produce the tuple $(x_1^j, \ldots, x_{k-1}^j, x_k)$.

Observe that for inputs given non-zero probability under distribution μ, $\textsc{Disj}_k(\vec{x}^j, x_k) = 1$ if and only if $j \in x_k$. It follows that the set $\{\textsc{Disj}_k(\vec{x}^j, x_k)\}_{j \in [n]}$ completely determines x_k. If the protocol P were always correct, then we could encode x_k by listing all the possible messages that could be sent by players $1, \ldots, k-1$ for any of the possible extensions $\vec{x}^j$ on the first j coordinates since these would be sufficient to determine the values of $\{\textsc{Disj}_k(\vec{x}^j, x_k)\}_{j \in [n]}$ and thus the bits of x_k. Although there are $n = m^{k-1}$ different extensions of x_k, for each player $1, \ldots, k-1$, given x_k there are only $m^{k-2} = n^{1-1/(k-1)}$ different messages possible since player i's message does not depend on the i-th coordinate. Thus the total number of bits required would be at most $(k-1)n^{1-1/(k-1)}C(P)$ which must be at least n since they are sufficient to encode x_k and we would obtain $C(P) \geq n^{1-1/(k-1)}/(k-1)$.

Since P has error at most ϵ this vector $\vec{v}$ of possible messages would instead only be sufficient to determine each bit of x_k with error at most ϵ under distribution μ. This implies that $\mathrm{H}_\mu(x_k \mid \vec{v}) \leq H_2(\epsilon)n$. Thus

$$\begin{aligned} n = \mathrm{H}_\mu(x_k) &\leq \mathrm{H}_\mu(\vec{v}) + \mathrm{H}_\mu(x_k \mid \vec{v}) \\ &\leq (k-1)n^{1-1/(k-1)}C(P) + H_2(\epsilon)n. \end{aligned}$$

Rearranging, we have $(k-1)n^{1-1/(k-1)}C(P) \geq (1 - H_2(\epsilon))n$ which yields the claimed bound. □

5.2 General k-party NOF Computation

We obtain lower bounds for general k-party NOF communication complexity as a simple consequence of Theorem 14 using a simulation of general protocols by simultaneous protocols.

Theorem 15. *For any $\epsilon < 1/2$, $R_\epsilon^k(\textsc{Disj}_k)$ is $\Omega(\frac{\log n}{k-1})$.*

Proof. Given an ϵ-error k-party NOF protocol P for $\textsc{Disj}_k$ of total complexity $C(P)$ define a simultaneous protocol P' for $\textsc{Disj}_k$ as follows. Each player sends a vector of length $2^{C(P)}$ that of all bits that the player would have sent in protocol P for every prefix of communications in which it is his turn to speak. Therefore by Theorem 14,

$$2^{R_\epsilon^k(\textsc{Disj}_k)} \geq (1 - H_2(\epsilon)) n^{1/(k-1)}/(k-1)$$

and thus

$$\begin{aligned} R_\epsilon^k(\textsc{Disj}_k) &\geq \log_2[(1 - H_2(\epsilon)) n^{1/(k-1)}/(k-1)] \\ &\geq \frac{\log_2 n}{k-1} - \log_2(\frac{k-1}{1 - H_2(\epsilon)}) \end{aligned}$$

which is $\Omega(\frac{\log n}{k-1})$. □

6 Discussion

Given that the best known upper bound for disjointness is $O(n)$ and the proximity of our $\Omega(\log n)$ lower bounds to the $\omega(\log^3 n)$ or $\omega(\log n (\log \log n)^2)$ lower bounds required for the proof complexity consequences in [8], it might seem that we have come most of the way to our goal. However there is still some way to go to understand the problem. For example it is not at all clear how one might extend the bound in Theorem 10 or even the one-way lower bound in [4] to 4 players. The problem is that it is not at all clear how

to prove any form of direct sum theorem even for one-way 3-party communication complexity. An impediment to extending our bounds to this case is the failure of the 3-party analogue of our method for Lemma 7 because, even for a cross-product distribution, the density of a 3-cylinder intersection is not determined by the densities of the cylinders involved in the intersection.

We have shown two different methods for deriving $\Omega(\log n)$ lower bounds on the general 3-party NOF complexity of disjointness. One reason to consider both methods is that the properties from which they are derived seem to be incomparable. The proof of Theorem 11 yields bounds on corruption for large 3-cylinder intersections that may be give useful insight into obtaining larger bounds. These bounds do not seem to follow from Theorem 15 but this has the advantage of a somewhat simpler proof and a result that applies more generally.

Finally, we note that independent of this work Klauck, Spalek, and de Wolf [21] derive similar bounds to Corollary 9 for 2-party quantum communication complexity using the polynomial method.

Acknowledgements

We would like to thank an anonymous reviewer for useful suggestions.

References

[1] N. Alon, Y. Matias, and M. Szegedy. The space complexity of approximating the frequency moments. *Journal of Computer and System Sciences*, 58(1):147–157, 1999.

[2] L. Babai, P. Frankl, and J. Simon. Complexity classes in communication complexity theory. In *27th Annual Symposium on Foundations of Computer Science*, pages 337–347, Toronto, Ontario, October 1986. IEEE.

[3] L. Babai, A. Gál, P. G. Kimmel, and S. V. Lokam. Communication complexity of simultaneous messages. *SIAM Journal on Computing*, 33(1):137–166, 2003.

[4] L. Babai, T. P. Hayes, and P. G. Kimmel. The cost of the missing bit: Communication complexity with help. *Combinatorica*, 21(4):455–488, 2001.

[5] L. Babai, N. Nisan, and M. Szegedy. Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *Journal of Computer and System Sciences*, 45(2):204–232, October 1992.

[6] Z. Bar-Yossef, T.S. Jayram, R. Kumar, and D. Sivakumar. Information theory methods in communication complexity. In *Proceedings Seventeenth Annual IEEE Conference on Computational Complexity*, pages 133–142, Montreal, PQ, Canada, May 2002.

[7] Z. Bar-Yossef, T.S. Jayram, R. Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. *Journal of Computer and System Sciences*, 68(4):702–732, 2004.

[8] P. Beame, T. Pitassi, and N. Segerlind. Polynomial threshold proofs and multiparty communication complexity. In *ICALP05*, 2005. Submitted.

[9] J.-Y. Cai. Lower bounds for constant depth circuits in the presence of help bits. *Information Processing Letters*, 36(2):79–83, 1990.

[10] A. Chakrabarti, S. Khot, and X. Sun. Near-optimal lower bounds on the multi-party communication complexity of set disjointness. In *Proceedings Eighteenth Annual IEEE Conference on Computational Complexity*, pages 107–117, Aarhus, Denmark, July 2003.

[11] A. Chakrabarti, Y. Shi, A. Wirth, and A.C-C. Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proceedings 42nd Annual Symposium on Foundations of Computer Science*, pages 270–278, Las Vegas, Nevada, October 2001. IEEE.

[12] A. K. Chandra, M. L. Furst, and R. J. Lipton. Multiparty protocols. In *Proceedings of the Fifteenth Annual ACM Symposium on Theory of Computing*, pages 94–99, Boston, MA, April 1983.

[13] F. R. K. Chung and P. Tetali. Communication complexity and quasi-randomness. *SIAM Journal on Discrete Mathematics*, 6(1):110–123, 1993.

[14] O. Goldreich, N. Nisan, and A. Wigderson. On Yao's XOR-lemma. Technical Report TR95-050, Electronic Colloquium in Computation Complexity, `http://www.eccc.uni-trier.de/eccc/`, 1995.

[15] R. Impagliazzo, R. Raz, and A. Wigderson. A direct product theorem. In *Proceedings, Structure in Complexity Theory, Ninth Annual Conference*, pages 88–97, Amsterdam, The Netherlands, June 1994. IEEE.

[16] R. Jain, J. Radhakrishnan, and P. Sen. A direct sum theorem in communication complexity via message compression. In J. C. M. Baeten, J. K. Lenstra, J. Parrow, and G. J. Woeginger, editors, *Automata, Languages, and Programming: 30th International Colloquium*, volume 2719 of *Lecture Notes in Computer*

Science, pages 300–315, Eindhoven, The Netherlands, July 2003. Springer-Verlag.

[17] B. Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. In *Proceedings, Structure in Complexity Theory, Second Annual Conference*, pages 41–49, Cornell University, Ithaca, NY, June 1987. IEEE.

[18] M. Karchmer, E. Kushilevitz, and N. Nisan. Fractional covers and communication complexity. In *Proceedings, Structure in Complexity Theory, Seventh Annual Conference*, pages 262–274, Boston, MA, June 1992. IEEE.

[19] M. Karchmer, R. Raz, and A. Wigderson. Superlogarithmic depth lower bounds via direct sum in communication complexity. *Computational Complexity*, 5:191–204, 1995.

[20] H. Klauck. Rectangle size bounds and threshold covers in communication complexity. In *Proceedings Eighteenth Annual IEEE Conference on Computational Complexity*, pages 118–134, Aarhus, Denmark, July 2003.

[21] H. Klauck, R. Spalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. In *Proceedings 45th Annual Symposium on Foundations of Computer Science*, Rome, Italy, October 2004. IEEE.

[22] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, Cambridge, England ; New York, 1997.

[23] N. Nisan, S. Rudich, and M. Saks. Products and help bits in decision trees. *SIAM Journal on Computing*, 28(3):1035–1050, 1999.

[24] N. Nisan and A. Wigderson. Rounds in communication complexity revisited. *SIAM Journal on Computing*, 22(1):211–219, 1993.

[25] R. Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(1):763–803, 1998.

[26] R. Raz. The BNS-Chung criterion for multi-party communication complexity. *Computational Complexity*, 9:113–122, 2000.

[27] Ran Raz and Avi Wigderson. Monotone circuits for matching require linear depth. *Journal of the ACM*, 39(3):736–744, July 1992.

[28] A. A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[29] M. E. Saks and X. Sun. Space lower bounds for distance approximation in the data stream model. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 360–369, Montreal, Quebec, Canada, May 2002.

[30] R. Shaltiel. Towards proving strong direct product theorems. In *Proceedings Sixteenth Annual IEEE Conference on Computational Complexity*, pages 107–117, Chicago, IL, June 2001.